

How Discovering New Data Will Be Key for Future BI

By Mike Lata

Originally Published on UBM Techweb's Big Data Republic

Business Intelligence (BI) involves the idea of transferring raw data (particularly unstructured) into meaningful insight to help improve a company's performance, enable executives to make future business decisions and forecast trends.

To gather the data needed for these tasks, a company should understand its market, customer base, potential customers (that can be targeted with marketing), and the types of data it has to work with. This, in turn, can lead to better business decisions and improved performance.

Data Discovery Techniques

Data discovery involves finding usable data for analysis that more often than not requires specific tools that makes it easy to explore data and find what is usable for your needs.

We [previously mentioned](#) five ways that data analysis and discovery techniques can bolster a company's performance. It is a growing trend to look for any competitive advantage a company can get in today's competitive markets and data discovery is one way of achieving this.

discovering relationships between data elements, metadata within data sets, data quality and consistency, data identification, campaign analysis, and other aspects of business intelligence within the big data sphere is what data discovery can offer your company.

An example of a tool that can be used to explore data and explore, analyze and discover usable data sets within the larger pool of data available, is [Spotfire](#). It is one of the many tools available on the market that can visualize the data discovery so that managers, contributors and executives can make sense of it through the use of graphs and other techniques.

[Visual data discovery](#) is a growing and important trend because it allows your company to manipulate, showcase and share various aspects of data as well as look for trends within it without relying on an IT team to make sense of it for you.

TIBCO has recently [announced](#) the release of the Spotfire data discovery and visualization platform, called TIBCO Spotfire Cloud, which may be worth taking a look at if you are looking for a data visualization solution for data discovery and business intelligence.

Being able to interact with the data is key with these tools because this will open up the door to see trends, productive business intelligence, and predictive analytics within your company.

Unstructured Data Discovery

Smart business intelligence involves analyzing a wide variety of data types for trends, and not just data that is easily put into rows and columns on spreadsheets. This other data that should be discovered and mined is none other than unstructured data.

This type of data may be hard to organize or put into traditional databases at first glance. It includes social media, video feeds, multimedia content, e-mails, word processing documents and presentations. Unstructured data is more open to interpretation than something traditionally stored in data warehouses.

Unstructured Data Sets and Hadoop

unstructured data is key to predictive analysis and business intelligence because it can open up new opportunities for insight. This is why companies should be using data discovery techniques using non-traditional sources of data like that of social media feeds.

Data that you may not even take into consideration can end up giving your company great insight after using proper analytics and data discovery techniques to make sense of it.

Unstructured data represents about 80% of all data out there. This data is out there for the taking, so to speak, and a framework such as Hadoop is great for this task.

Using Hadoop a company can input the unstructured data into its file system (HDFS) and scale it with multiple nodes of servers running in a cluster in parallel computing using MapReduce. This means it is highly flexible and scalable to very large data sets coming in constant streams for batch processing and the processing can be spread out over thousands of servers working together for the task.

Data Discovery Integration with Hadoop

Hadoop has a lot of benefits that can be turned into business intelligence and predictive analytics. It is highly flexible with the type of data its HDFS can work with and offers a great way to do parallel computation for efficient processing of the data.

However, Hadoop requires a connector that will let your company add data discovery techniques onto its platform. This connector will integrate your data visually and efficiently into operational data.

A connector you should consider utilizing for this task is the [Spotfire in-memory analytics platform](#) that is part of Spotfire 4.5. Its strengths include self-service data discovery, fastest to actionable insight, visibility into the unknown, and universal adaptability.

It is highly flexible to be deployed within a wide variety of organizations and you can see visual solutions with it across a wide variety of devices, such as tablets and smartphones. Its connection to Hadoop, however, is what truly makes it stand out.

“The new Spotfire 4.5 adds a connector for Apache Hadoop, which will let users add big data to data mashups, data discovery and analytics tasks they now do with data from Oracle, SAP, Salesforce.com and unstructured data,” according to [Integration Developer News](#).

This is a great way to tap into unstructured data sources running from Hadoop and make sense of it all visually.

Data Discovery and the Internet of Things

Besides simply reporting what is out there, new techniques are starting to open up to find new sources of data that can be turned into unstructured data sets and later BI. Data discovery is a field with many opportunities as data come from a wide range of sources.

Sensors, for instance, can be used inside cars as data discovery techniques. Objects that fall into [the Internet of Things](#) framework can also be used such as refrigerators with WiFi access. In-car location tracking can be another data discovery technique for further analysis.

As you see, there is a lot of data that has yet to be discovered or taken for consideration when it comes to analytics and BI. The future is bright and more companies are starting to find ways to input this data in databases or frameworks such as Hadoop that can make sense of it all, but it needs an efficient connector and integration such as Spotfire 4.5.